

ABSTRACT

The rough set theory-based reduct algorithm is an effective approach for filtering irrelevant or redundant attributes from datasets, especially in the context of large-scale data classification. By reducing the number of attributes without sacrificing information accuracy, reduct plays a crucial role in improving the efficiency and performance of classification models. In this study, the reduct algorithm is combined with a Genetic Algorithm (GA)-based heuristic method to perform optimal feature selection. Genetic Algorithm (GA) is used to explore the global solution space and find the best combination of attributes that support efficient multi-class classification. This approach is implemented in the case of property price classification in the real estate sector's e-commerce ecosystem, which is now rapidly growing along with digitalization. The dataset used is public property sales data (Real Estate Sales 2001–2022 GL) from the State of Connecticut, which is then analyzed in the context of property price classification. The results show that the combination of attribute reduction and Genetic Algorithm (GA)-based feature selection is consistently able to select significant attributes, such as assessed value, sales ratio, and property type. Model evaluation yielded a classification accuracy of 96%–98% with an average macro f1-score of 0.94–0.98, and a classification time of less than 0.13–0.22 seconds. However, the feature selection process using the Genetic Algorithm (GA) takes a total time of approximately 4–15 minutes, which can be an efficiency constraint on large datasets. Therefore, this approach demonstrates effectiveness in optimizing multi-class classification on real estate data and can be applied to support data-driven decision-making in the property e-commerce ecosystem.

Keywords: *E-commerce*, property, price classification, *reduct*, *rough set*, *Genetic Algorithm*, heuristic, feature selection.